USAARL Report No. 99-15

Evaluating Performance Effects of a Medication (Dexedrine®) in the Simulator Versus Aircraft Environment

By

John A. Caldwell
Kristi A. Roberts
Heber D. Jones

Aircrew Health and Performance Division

19990517 076

April 1999

DTIC QUALITY INSPECTED 4

Approved for public release, distribution unlimited.

U.S. Army Aeromedical Research Laboratory
Fort Rucker, Alabama 36362-0577

Notice

Qualified requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Cameron Station, Alexandria, Virginia 22314. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Human use

Human subjects participated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR 70-25 and USAMRMC Reg 70-25 on Use of Volunteers in Research.

Reviewed:

MORRIS R. LATTIMORE, JR.
Colonel, MS
Director, Aircrew Health &
Performance Division

Released for publication:

JOHN A. CALDWELL, Ph.D.
Chairman, Scientific Review
Committee

CHERRY L. GAFFNEY
Colonel, MC, SFS
Commanding

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release, distribution unlimited

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
USAARL Report No. 99-15

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
U.S. Army Aeromedical Research Laboratory

6b. OFFICE SYMBOL (If applicable)
MCMR-UAS

7a. NAME OF MONITORING ORGANIZATION
U.S. Army Medical Research and Materiel Command

6c. ADDRESS (City, State, and ZIP Code)
P.O. Box 620577
Fort Rucker, AL 36362-0577

7b. ADDRESS (City, State, and ZIP Code)
Fort Detrick
Frederick, MD 21702-5012

8a. NAME OF FUNDING / SPONSORING ORGANIZATION

8b. OFFICE SYMBOL (If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
62787A	30162787A879	OC	DA336185

11. TITLE (Include Security Classification) (U) Evaluating Performance Effects of a Medication (Dexedrine®) in the Simulator Versus Aircraft Environment

12. PERSONAL AUTHOR(S)
J.A. Caldwell; K.A. Roberts; and H.D. Jones

13a. TYPE OF REPORT	13b. TIME COVERED FROM	TO	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final			1999 April	19

16. SUPPLEMENTAL NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
06	04	
06	15	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Flight performance, simulator, aviation, MATB, UH-60, dextroamphetamine

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A quasi-experimental approach assessed the pilot performance comparability of simulator versus in-flight results. Flight data from three sleep-deprivation studies were pooled. Twenty aviators were included, 10 who flew a UH-60 helicopter simulator and 10 who flew a UH-60A aircraft under the influence of Dexedrine or placebo during 40 hours of continuous wakefulness. Performance on straight and levels, right and left turns, climbs and descents, and a left-descending turn (assessed at 0100, 0500, 0900, 1300, and 1700) tended to correspond in the simulator and aircraft. Generally, performance under Dexedrine was better than under placebo. However, only half of the maneuvers showed consistent, statistically significant stimulant/fatigue effects in both flight platforms. Measurement sensitivity was lower in the aircraft, likely because of error variance due to environmental influences (weather, temperature, and turbulence) and other factors (radio traffic and anxieties about safety). Thus, actual in-flight studies, while desirable in terms of face-validity, underestimate the impact of stressors such as fatigue on pilots.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Chief, Science Support Center

22b. TELEPHONE (Include Area Code)
(334) 255-6907

22c. OFFICE SYMBOL
MCMR-UAX-SS

DD Form 1473, JUN 86 *Previous editions are obsolete.*

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

Acknowledgments

The authors wish to thank Dr. J. Lynn Caldwell for scoring the simulator data. In addition, the team effort of a variety of personnel within the U.S. Army Aeromedical Research Laboratory and all of the volunteer helicopter pilots who made this study possible, is sincerely appreciated.

Table of contents

Page

Background 1

Methods 2
Subjects 2
Apparatus 3
Drug dosing 3
UH-60 flight simulator 3
UH-60 helicopter 3

Procedure 3
Flight evaluations 4
Testing schedule 5

Results 5
Flight performance data 5
Straight and levels 6
Left standard-rate turns 6
Climbs 8
Right standard-rate turns 9
Descents 9
Left descending turn 10

Discussion 10

References 13

List of Tables

1. Flight profile 4
2. Scoring bands for flight performance data 5
3. Means for SL iterations under placebo and Dexedrine 6
4. Means for LSRTs under placebo and Dexedrine 7
5. Means for climbs in simulator versus aircraft flights 8
6. Means for RSRT iterations in simulator versus aircraft 9

List of Figures

1. Effect of drug and session on LSRT scores 7
2. Effects of drug and session on scores from the straight climbs 8
3. Effects of drug by session on scores from the straight descents 10

Background

Accurate measurement of pilot performance has been of interest to the aviation community for years (Rehmann, 1982). The implementation of new operational procedures, fielding of various pharmacological interventions, understanding of stressor effects, and improvement of training and tactical operations all rely upon sensitive and reliable methods of evaluating aviator performance. Due to technological advances, it is now possible to examine piloting skill via computerized flight scoring systems both in simulators and aircraft. However, while some studies have been conducted under actual flight conditions (Billings, et al., 1968; Billings, Gerke, and Wick, 1975; Caldwell and Caldwell, 1997; and Caldwell, Stephens, and Carter, 1992) the majority have relied on simulators (Caldwell, Caldwell, and Crowley, 1996; Caldwell et al., 1995; Caldwell et al., 1996; Dellinger, Taylor, and Porges, 1987; Henry et al., 1974; Simmons et al., 1989; Stephens et al., 1992).

Simulator studies are attractive because of low relative cost, greater accessibility, optimal experimental control, and improved safety relative to in-flight investigations. Simulations have contributed much toward the understanding of aviation-related problems and the effects of stressors or interventions. One category of pilot-performance study that has benefited from flight simulation is the area of drug research. Studies on several compounds have been published over the past 10 years, and each has yielded valuable information for the operational environment. For example, studies on the chemical defense compound, atropine sulfate, have characterized the drug's effects and eased concerns about the fielding of this medication. Dellinger, Taylor and Porges (1987) and Simmons et al. (1989) showed that an atropine injection did not preclude an aviator's safe return to base despite the presence of performance decrements with large doses. Other studies have shown that the antihistamine terfenadine is safe for aviators because it does not degrade performance (Stephens et al., 1992); the hypnotic triazolam, while effective, is of limited use in pilots because of its potential side effects (Caldwell et al., 1996); and the stimulant dextroamphetamine is efficacious for maintaining aviator performance during sleep loss (Caldwell et al., 1995 and Caldwell et al., 1996).

These findings presumably apply to the aircraft environment but there is limited empirical evidence on this point. Typically, the option of conducting in-flight studies has been abandoned in favor of working in a simulator because of feasibility and safety factors. Simulations possess a high degree of face validity, and convenience factors make them a highly attractive alternative to the aircraft. However, since few investigators have the time or resources necessary to perform simulator versus in-flight comparability studies, it is unclear how well findings from one situation will actually generalize to the other.

The few comparability studies that do exist suggest that simulations are more sensitive than aircraft studies to performance changes. Caldwell and Jones (1990) compared helicopter in-flight data to helicopter simulator data collected as part of the atropine work mentioned above. They concluded the simulator offered the most sensitivity to drug effects, especially when a small or moderately impairing dose (2 mg) was used. Billings, Gerke and Wick (1975) reported similar findings (simulator more sensitive than aircraft) in their work with secobarbitol, particularly when low doses of the drug (100 mg) were tested. They concluded simulators were

useful for sensitive, inexpensive investigations of stressors and pilot performance; however, they advised caution when extrapolating from the simulator to the aircraft because of the differences in pilot arousal levels from one situation to the other.

How much of a difference really exists between the results of laboratory simulations and actual in-flight investigations, and is this difference a genuine cause for concern? Based on the limited number of available comparison studies, it is difficult to answer these questions. On the one hand, it is possible that the only good reasons for conducting simulations are cost and safety related, and that all other things being equal, an in-flight investigation is the most desirable alternative. On the other hand, perhaps the increased sensitivity in the simulator environment provides information that is totally obscured in the more realistic, but more variable, in-flight domain.

The present paper attempts to address these issues by comparing simulator and aircraft data collected during three studies on the effects of dextroamphetamine in sleep-deprived pilots. The first was a study by Caldwell et al. (1995) in which six male aviators were kept awake for 40 continuous hours while they flew a helicopter simulator and completed other evaluations. During the last half of one period, the subjects were administered 10-mg doses of Dexedrine (at 0000, 0400, and 0800), and during the final hours of the other period, the subjects were given placebo. Dexedrine improved composite measures of flight performance on four out of six sets of maneuvers, with the most notable benefits occurring at 0500 and 0900 when the fatigue was most severe. These results were confirmed by Caldwell et al. (1996) in a systematic replication of the 1995 study . In this case, six females were used as subjects. Once again, the simulator flight data showed the majority of maneuvers were better after Dexedrine than placebo, and often, there were drug by time-of-day effects. It was concluded that Dexedrine was a viable fatigue countermeasure for sleep-deprived pilots. However, since both tests were conducted in a simulator, an in-flight study was felt to be necessary before definitive conclusions were possible. Thus, in 1997, a systematic replication was performed in a specially-instrumented UH-60 helicopter (Caldwell and Caldwell, 1997). Results again indicated improved performance with Dexedrine on many maneuvers; however, the impact of the drug was not as robust as in the earlier simulator investigations. In fact, many of the drug-by-time effects (seen in the simulator) were less pronounced or absent in the aircraft, and composite scores proved inadequate to detect many of the in-flight drug effects (root mean square errors of control parameters were used instead).

Both these simulator and in-flight investigations offered evidence of the efficacy of dextroamphetamine, but there were differences. To examine the extent of these differences, the present investigation was performed.

Methods

Subjects

Two groups of subjects were compared. Ten UH-60 pilots (mean age of 28.3 years) were selected from among the 12 pilots who contributed data in the simulator studies. The final group

consisted of 5 males and 5 females who were combined to create the "simulator group." All 10 UH-60 pilots (mean age of 31.9 years) who contributed to the in-flight study were used in the "in-flight group." These were all males. It was felt acceptable to combine the male and female subjects in the simulator group based on an analysis which showed no differences between the genders in flight performance. Female volunteers were screened for pregnancy prior to admission. The average amount of flight experience for the participants in the simulator study was 1,003 hours (ranging from 140-3,400 hours) and the average flight experience for the participants in the in-flight study was 1,278 hours (ranging from 540-3,100 hours). The approximate average weights for participants in the two groups averaged 150 and 155 pounds, respectively.

Apparatus

Drug dosing

At each dose, subjects received two orange capsules (placebo or Dexedrine®) with 8 ounces of orange juice. Placebo capsules were filled with lactose, and each of the Dexedrine® capsules contained one, 5-mg Dexedrine® tablet. Dosages were not adjusted according to the body weights of subjects because similar adjustments would not be performed under the field conditions which this research was designed to simulate.

UH-60 flight simulator

Simulator flights were conducted in a UH-60 simulator with a 6-degree-of-freedom motion base and a full-visual cockpit in which the visual display was set for daytime flight. Flight data (heading, airspeed, altitude, etc.) were acquired by computer and converted to composite flight scores using specialized routines (Jones and Higdon, 1991).

UH-60 helicopter

In-flight evaluations were conducted in a specially-instrumented Sikorsky JUH-60A helicopter. Both day and night flights were conducted under unaided conditions (night vision goggles were not used at night). Flight data were recorded with a locally-manufactured, computerized flight monitoring package referred to as the Aeromedical Instrumentation System (AIS). Data from the AIS were converted to composite flight scores using the software routines mentioned above.

Procedure

Each subject completed several flights under Dexedrine® and placebo. The dose-administration schedule was fully counterbalanced and double blind.

Flight evaluations

Flight performance evaluations required subjects to perform a variety of instrument flight maneuvers arranged in a standardized upper-airwork profile. These maneuvers required reliance on aircraft/simulator flight instruments rather than external visual cues. In the simulator, subjects began by performing hovers and low-level navigation tasks followed by instrument maneuvers and a formation flight, but only the instrument maneuvers are examined here. In the aircraft, subjects began by flying the aircraft to a safe maneuvering area prior to performing the same instrument maneuvers which were used in the simulator flight (after the hovers and navigation). The last 1000-foot descent was deleted from the aircraft flight profile for safety reasons (thus, the lowest altitude of any flight maneuver was 1700 feet above the ground).

Maneuvers were flown in the same order each time (see table 1). The first group of these was flown with the automatic flight control system (AFCS) trim engaged (the normal mode when flying the UH-60), and the second group was flown with the AFCS trim turned off.

Table 1.
Flight profile.

Manueuver	AFCS On/Off
Straight and level number 1	On
Left standard-rate turn number 1	On
Straight and level number 2	On
Climb number 1	On
Right standard-rate turn number 1	On
Straight and level number 3	On
Right standard-rate turn number 2	On
Climb number 2	On
Descent number 1	Off
Left descending turn	Off
Descent number 2	Off
Left standard-rate turn number 3	Off
Straight and level number 4	Off
Right standard-rate turn number 3	Off

The AFCS trim system enhances the stability of the aircraft/simulator, and when the AFCS is turned off (to simulate a system failure), accurate flight control becomes much more difficult, increasing the pilot's workload.

During each maneuver, the subject was required to maintain control over specific flight parameters (i.e., heading, altitude, etc.) which varied across maneuvers. For instance, heading control was evaluated during straight-and-level flight, but not turns. Scores which reflected how well the subject flew each maneuver were calculated in two steps. First, the control scores for the parameters relevant to each maneuver were determined using the limits presented in table 2. Thus, if a subject never deviated from the assigned heading by more than 1 degree, a score of

100 resulted, whereas larger deviations produced lower scores. Second, the scores from each parameter were averaged into a single composite score. Thus, if a subject scored 100 on heading, 85 on altitude, and 90 on airspeed, a composite score of 91.7 would have resulted. Composite scores were not collapsed across all of the maneuvers in each flight because of the differences in the parameters which made up the scores in each. To ensure there were no large "offset errors" attributable to subjects failing to establish correct headings, altitudes, airspeeds, etc., at the outset, the volunteers were required to attain correct flight parameters before the beginning of each maneuver.

Table 2.
Scoring bands for flight performance data.

Maximum deviations for scores of:						
Measure (units)	100.0	80.0	60.0	40.0	20.0	0
Heading (degrees)	1.0	2.0	4.0	8.0	16.0 >	16.0
Altitude (feet)	8.8	17.5	35.0	70.0	140.0 >	140.0
Airspeed (knots)	1.3	2.5	5.0	10.0	20.0 >	20.0
Slip (ball widths)	0.0	0.1	0.2	0.4	0.8 >	0.8
Roll (degrees)	0.8	1.5	3.0	6.0	12.0 >	12.0
Vertical Speed (feet/m)	10.0	20.0	40.0	80.0	160.0 >	160.0
Turn Rate (degrees/s)	0.3	0.5	1.0	2.0	4.0 >	4.0

Testing schedule

Subjects arrived at the laboratory at 1800 on Sunday when the study was explained, informed consent was obtained, and a medical evaluation was conducted. Subjects with past psychiatric or cardiac disorder, a history of sleep disturbances, or any current significant illness would have been rejected, but none of these problems were found. On Monday morning, the aviator completed three training flights (at 0900, 1300, and 1700) before retiring at 2300 hours. On Tuesday, there were three control-day flights (at 0900, 1300, and 1700), but at the end of the day, sleep was not permitted. On Wednesday at 0000 the first drug/placebo dose was administered, followed by subsequent doses at 0400 and 0800. Flights occurred at 0100, 0500, 0900, 1300, and 1700. On Thursday, after 8 hours of recovery sleep, the subject repeated the same schedule as was used on Tuesday. On Friday, testing continued with drug/placebo doses at 0000, 0400, and 0800, and flights were conducted at the same times as those on Wednesday. The participant retired at 2300 on Friday and was released after awakening at 0700 on Saturday morning.

Results

Flight performance data

Analysis of variance (ANOVA) was used to analyze scores for each maneuver. The between-subjects factor was group (simulator, aircraft) and the within-subjects factors were drug (placebo, Dexedrine) and session (0100, 0500, 0900, 1300, 1700). For maneuvers flown more than once

during the profile, a third factor, iteration (i.e., turn 1 and turn 2), was added. Significant interactions and main effects were followed by analysis of simple effects and/or pairwise contrasts. Huynh-Feldt adjusted degrees of freedom were used in the event of violations of the compound symmetry assumption. Only the effects involving group or drug are discussed below.

Straight and levels

The ANOVA for the four iterations of straight and levels (SLs 1-4) indicated a drug-by-iteration-by-group interaction ($F(3, 54)=5.09$, $p=.0036$) due to a drug-by-iteration effect in the simulator but not the aircraft. In the simulator, performance under placebo was lower than under Dexedrine during both SL2 and SL4, while at SL1 and SL3 there was no drug-related difference (see table 3).

Table 3.
Means for SL iterations under placebo and Dexedrine.

Group	Drug	SL1	SL2	SL3	SL4
Simulator	Pbo	89.7	84.4	82.6	74.6
Simulator	Dex	91.7	87.7	85.7	84.3
Aircraft	Pbo	73.0	68.0	66.5	71.1
Aircraft	Dex	74.5	71.1	69.0	73.7

There was an iteration-by-group interaction ($F(2.78, 50.02)=16.31$, $p<.0001$). In both groups, higher scores occurred during SL1 than in SL2 or SL3, but only in the simulator was there a further drop at SL4. In the aircraft, scores during this last SL were slightly higher than scores at SL2 and SL3. A drug-by-iteration interaction ($F(3, 54)=6.57$, $p=.0007$) was due to differences across the SLs under placebo versus Dexedrine. There was no drug effect in SL1, but Dexedrine produced moderately better performance than placebo in SL2 and SL3, and much better performance in SL4.

There were main effects on group ($F(1,18)=53.70$, $p<.0001$)and drug ($F(1,18)=16.17$, $p=.0008$). Scores were higher in the simulator than in the aircraft (85.08 versus 70.86) and higher under Dexedrine than placebo (79.71 versus 76.23).

Left standard-rate turns

The ANOVA for the left standard-rate turns (LSRT1, LSRT2) indicated a drug-by-iteration-by-group interaction for scores ($F(1,18)=5.12$, $p=.0363$) which analysis of simple effects indicated was due to a drug-by-iteration effect in the simulator, but not in the aircraft. In the simulator, performance under placebo was worse than performance under Dexedrine at LSRT2 while there was no difference at LSRT1 (see table 4).

Table 4.
Means for LSRTs under placebo and Dexedrine.

Group	Drug	LSRT1	LSRT2
Simulator	Pbo	78.6	59.1
Simulator	Dex	80.2	66.7
Aircraft	Pbo	62.2	60.5
Aircraft	Dex	63.6	62.3

There was a session-by-group interaction ($F(4, 72)=2.45$, $p=.0540$) due to differences across the testing times in the aircraft but not in the simulator. In the aircraft, there were higher scores at 0100 than at 0500 or 1300, and higher scores at 0900 than at 1300. In addition, there was a slight recovery in performance from 1300 to 1700 (the means for each flight from 0100-1700 in the aircraft were 63.9, 62.0, 62.4, 59.6, and 62.8, respectively). There was an iteration-by-group interaction ($F(1,18)=50.97$, $p<.0001$) due to better performance during LSRT1 than during LSRT2 in the simulator, but not in the aircraft. There was a drug-by-session interaction ($F(3.57,64.27)=3.64$, $p=.0125$) due to differences among testing times under placebo but not Dexedrine. Scores under placebo dropped from 0100 to 0500, 0900, and 1300, after which there was a recovery from 0900 to 1700 (see figure 1). There was a drug-by-iteration interaction ($F(1,18)=7.06$, $p=.0161$) due to a significant drug effect in the second, but not the first LSRT (higher scores under Dexedrine than placebo in LSRT2).

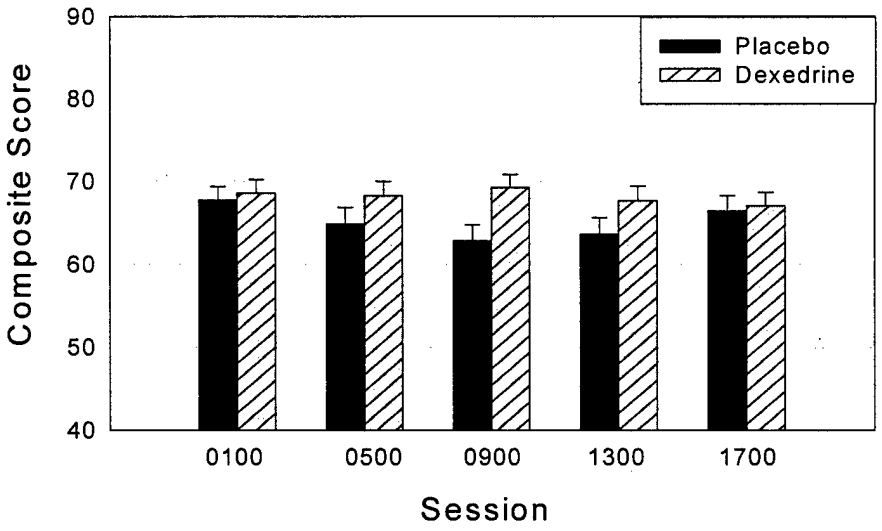

Figure 1. Effects of drug and session on LSRT scores.

There was a main effect on group ($F(1,18)=10.76$, $p=.0042$) because of better overall performance in the simulator than the aircraft (71.16 versus 62.13), and there was a main effect on drug ($F(1,18)=10.64$, $p=.0043$) due to higher composite scores under Dexedrine than placebo (68.18 versus 65.12).

Climbs

The ANOVA for the two climbs (Climb1, Climb2) indicated a drug-by-session-by-iteration interaction ($F(3.38,60.85)=2.69$, $p=.0480$) which analysis of simple effects indicated was due to a drug-by-session interaction at Climb2, but not Climb1. Analysis showed that although there were session differences both under placebo and Dexedrine ($p<.01$), the pattern was different. Under placebo, performance was better at 0100 than at 0900 or 1300, and performance was better at 0500 than at 0900. Under Dexedrine, there were no differences among the first three sessions, but performance at both 0500 and 0900 was better than performance at 1300, and performance at 0900 also was better than at 1700 (see figure 2)

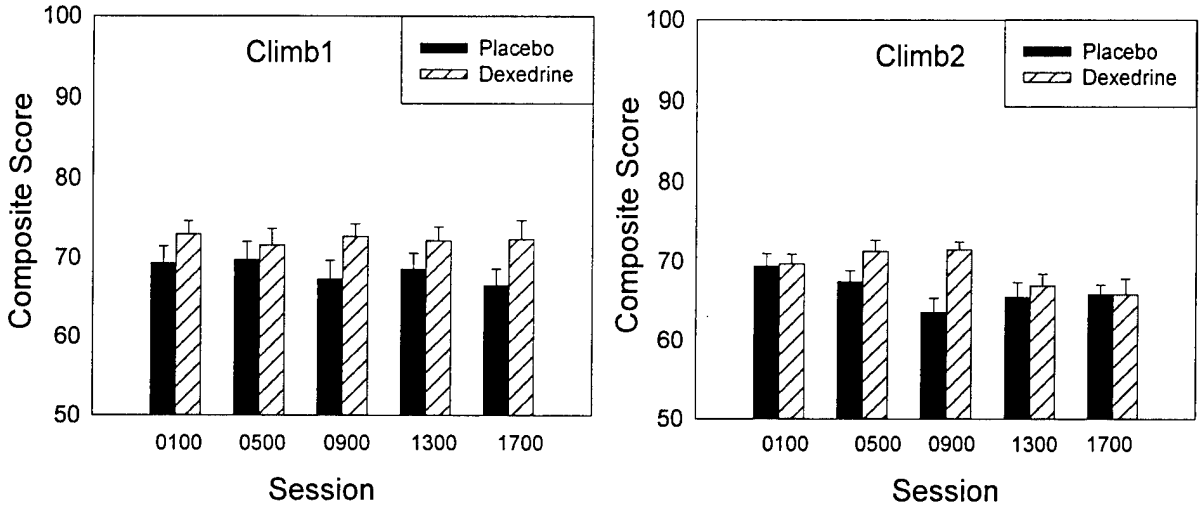

Figure 2. Effects of drug and session on scores from the straight climbs.

There was a session-by-group interaction ($F(3.65,65.73)=3.88$, $p=.0086$) due to differences across the testing times in the aircraft, but not in the simulator. In the aircraft, performance was better at 0100, 0500, and 0900 than at 1300, and performance was better at 0100 than at 1700 (see table 5).

Table 5.
Means for climbs in simulator versus aircraft flights.

Group	0100	0500	0900	1300	1700
Simulator	72.9	71.9	70.5	73.4	71.5
Aircraft	67.8	68.2	67.1	63.1	63.7

An iteration-by-group interaction ($F(1,18)=13.07$, $p=.0020$) occurred due to higher scores during Climb1 than Climb2 in the simulator, but not the aircraft. A group main effect ($F(1,18)=3.11$, $p=.0020$) was found due to higher scores in the simulator than the aircraft (72.05

versus 65.98), and a drug main effect ($F(1,18)=14.18$, $p=.0014$) occurred because of better performance under Dexedrine than placebo (70.73 versus 67.30).

Right standard-rate turns

The ANOVA for the three right standard-rate turns (RSRT1, RSRT2, RSRT3) indicated there was no 3-way interaction. However, there was a drug-by-group interaction ($F(1,18)=8.84$, $p=.0082$) due to better performance under Dexedrine than placebo in the simulator, but not in the aircraft. There also was an iteration-by-group interaction ($F(2,36)=14.76$, $p<.0001$) which was again because of an effect only in the simulator. In the simulator, scores in RSRT2 were better than those in the other two iterations, and RSRT1 was better than RSRT3 (see table 6).

Table 6.
Means for RSRT iterations in simulator versus aircraft.

Group	RSRT1	RSRT2	RSRT3
Simulator	74.4	78.1	69.8
Aircraft	60.2	60.7	60.2

There was a drug-by-iteration interaction ($F(2,36)=3.51$, $p=.0404$) because of differences in the drug effects across iterations. Dexedrine was associated with better performance than placebo in all three RSRTs; however, the difference was larger in RSRT3 than in RSRT1 and RSRT2.

In addition to these interactions, there was a group effect ($F(1,18)=23.30$, $p=.0001$) attributable to better performance in the simulator than in the aircraft (74.10 versus 60.39); and a drug effect ($F(1,18)=19.88$, $p=.0003$) due to higher scores under Dexedrine than placebo (69.17 versus 65.32).

Descents

The ANOVA for the two iterations of descents indicated a drug-by-group interaction ($F(1,18)=7.74$, $p=.0123$). Although there were Dexedrine-related improvements in both the simulator and the aircraft, it was most pronounced in the simulator. A drug-by-session interaction ($F(4,72)=3.79$, $p=.0074$) was due to differences across the testing times under placebo but not Dexedrine. Scores under placebo dropped substantially from the 0100 flight in comparison to the remaining flights. In addition, scores at 0500 were higher than those at 0900 (see figure 3).

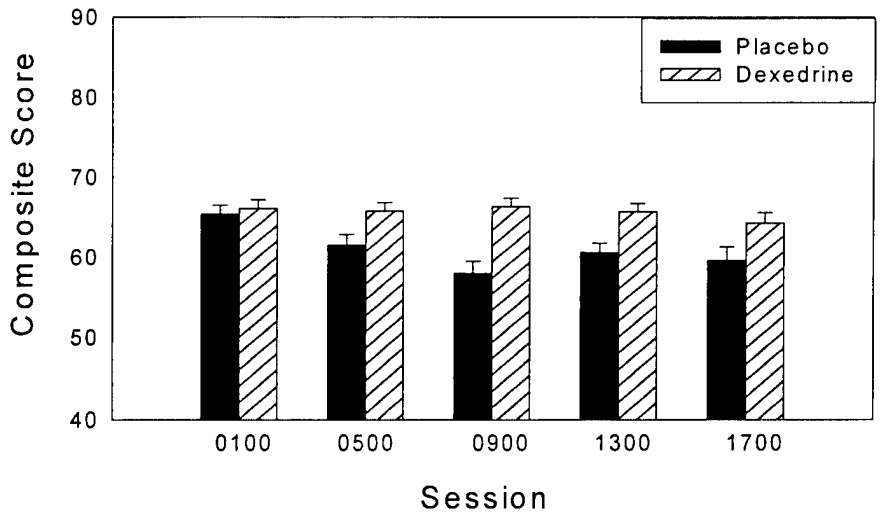

Figure 3. Effects of drug and session on scores from the straight descents.

Lastly in the descents, there was a drug main effect ($F(1,18)=35.84$, $p<.0001$) due to higher scores under Dexedrine than placebo. The means were 65.66 versus 61.03.

Left descending turn

The ANOVA for the left descending turn showed a drug-by-group interaction ($F(1,18)=5.38$, $p=.0323$) due to higher scores under Dexedrine than placebo in the simulator, but not in the aircraft. Also, there was a drug main effect ($F(1,18)=13.45$, $p=.0018$) because of better performance under Dexedrine than placebo (55.44 versus 51.61). There were no overall differences on the grouping factor (simulator versus aircraft).

Discussion

Both the simulator and aircraft data reported here were collected with virtually identical protocols which involved the same drug doses, test schedules, and experimental procedures. The overall findings from both indicated Dexedrine was associated with better performance than placebo in sleep deprived pilots. Statistically significant drug main effects occurred on every maneuver--a finding consistent with previous reports which have shown Dexedrine effectively attenuates the performance declines associated with sleep loss (Caldwell and Caldwell, 1997; Caldwell et al., 1995; and Caldwell et al., 1996). However, interpretation of some of the drug effects was complicated by the presence of interactions suggesting differences between the simulator and in-flight testing.

Differences between drug-related effects in the simulator versus the aircraft were seen in five of the six maneuvers. In four cases, drug effects were found in the simulator which did not attain statistical significance in the aircraft, and in one case, the observed drug effects were more pronounced in the simulator than in the aircraft (although differences did occur in both). These

findings suggest a higher degree of measurement sensitivity in the simulator environment which, in some situations, could lead to research conclusions that may not generalize in a straightforward manner to the actual flight environment. This finding supports Billings, Gerke, and Wick (1975) and Caldwell and Jones (1990) who concluded drug effects were more consistent and orderly in simulator than in aircraft tests.

There are a number of possible reasons for differences in the two situations, but the first and most probable is that weather turbulence, which creates large, frequent, and random flight-path deviations, is omnipresent in the aircraft and totally absent in the simulator. This has the net effect of reducing the accuracy of in-flight performance by causing the pilot to constantly correct for deviations which are unpredictably induced by wind gusts or thermal air currents. That this was an issue in the present study was evidenced by the presence of group main effects in two-thirds of the maneuvers. Wind turbulence accentuates statistical error variance to the point where only the most robust drug (or other) effects are large enough to outweigh random sources of performance variability. Thus, while there were consistent tendencies for performance to have been better under Dexedrine than placebo throughout all of the data, the differences sometimes were not large enough (in relationship to other sources of variance) to attain statistical significance.

The second explanation is that arousal levels in the actual flight environment may have been substantially higher than those in the simulator. This arousal difference may have increased performance capacity under the placebo condition to the point where some of the effects of sleep deprivation may have been overcome by anxiety alone. Thus, although Dexedrine improved alertness both in the simulator and the aircraft, the improvement relative to the no-drug condition tended to be smaller under actual flight conditions.

Other possible explanations for the differences between simulator and in-flight results include: environmental changes (in constrast to the simulator study, it was impossible to maintain a constant temperature and illumination level from one test period to another in the aircraft); differences in instructor pilots (flight-hour or crew-rest restrictions forced the use of different safety pilots in the aircraft but not in the simulator); and timing fluctuations (the simulator sessions always began precisely on time, whereas air traffic considerations sometimes introduced delays under actual flight conditions). Also, the results may have been influenced by the fact that there were differences in the pilot experience levels between the two groups. Subjects in the in-flight group had almost 300 hours more flight time than those in the simulator group. Thus, in-flight participants may have been better equipped to deal with the effects of fatigue under the placebo condition, and this may have minimized the apparent benefits from Dexedrine. Any of these factors could have decreased the sensitivity of the in-flight study.

Of course it is no surprise that a tightly controlled laboratory experiment would have yielded results different to those obtained in the real world, but it is interesting to note the extent to which these differences might have clouded the conclusions from at least one of our investigations. Based on the in-flight results alone, the efficacy of Dexedrine for sustaining flight performance during sleep loss would have been underestimated, and this would have been inconsistent with the robust improvements observed in the appearance, behavior, mood, and

physiological arousal levels of the research subjects. It is interesting to note that in both the in-flight and simulator investigations, staff members, safety pilots, and the volunteers easily differentiated between the Dexedrine and placebo conditions before the blinding procedure was removed. The fact that these robust effects did not manifest themselves more substantially in actual in-flight performance is unfortunate; however, because of the simulator investigations, we were able to attribute this difference to methodology rather than the intervention itself. In this case, the simulator made clear the beneficial effects of an interevention that otherwise might have been overlooked. If the drug under consideration had been one that impaired rather than improved performance, a similar problem would have arisen (lack of sensitivity in the aircraft), but the consequences could have been more problematic since it might have been concluded that this hypothetical drug was safe for flight operations when in fact the opposite was true. Because of this, it is recommended that testing be conducted first in the laboratory to gain a thorough understanding of the effects of any stressor or countermeasure on aviator performance, before conducting in-flight evaluations to "prove the concept" in the real world.

References

Billings, C. E., Eggspuehler, J. J., Gerke, R. J., Chase, R. C. 1968. Studies of pilot performance: II. Evaluation of performance during low altitude flights in helicopters. Aerospace Medicine, 39, 19-31.

Billings, C. E., Gerke, R. J., and Wick, R. L. 1975. Comparisons of pilot performance in simulated and actual flight. Aviation, Space and Environmental Medicine. 46, 304-308.

Caldwell, J. A. and Caldwell, J. L. 1997. An in-flight investigation of the efficacy of dextroamphetamine for sustaining helicopter pilot performance. Aviation, Space, and Environmental Medicine. 68(12), 1073-1080.

Caldwell, J. A., Caldwell, J. L., Crowley, J. S., and Jones, H. D. 1995. Sustaining helicopter pilot performance with Dexedrine during periods of sleep deprivation. Aviation, Space, and Environmental Medicine. 66(10), 930-937.

Caldwell, J. A., Caldwell, J. L., and Crowley, J. S. 1996. Sustaining female helicopter pilot performance with Dexedrine®during sustained operations. International Journal of Aviation Psychology. 7(1), 15-36.

Caldwell, J.A., and Jones, H. D. 1990. A comparison of computerized measurements of helicopter pilot performance with atropine sulfate during actual and simulated flight. AGARD Conference Proceedings No. 490, pp 5-1 to 5-10. North Atlantic Treaty Organization. Neuilly Sur Seine, France.

Caldwell, J. A., Stephens, R. L., and Carter, D. J. 1992. The effects of 2 mg and 4 mg atropine sulfate on the performance of U.S. Army helicopter pilots. Aviation, Space, and Environmental Medicine. 63(10), 857-864.

Caldwell, J. L., Caldwell, J. A., Sprenger, W. D., Jones, H. D., Woodrum, L., Colon, J. A., and Barto, J. 1996. Effects of triazolam on pilot performance. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 96-11.

Dellinger, J. A., Taylor, H. L., and Porges, S. W. 1987. Atropine sulfate effects on aviator performance and on respiratory-heart period interactions. Aviation, Space, and Environmental Medicine. 58, 333-338.

Henry, P. H., Flueck, J. A., Samford, J. F., Keiser, H. N., McNee, R. C., Walter, W. H., Webster, K. H., Hartman, B. O., and Lancaster, M. C. 1974. Assessment of performance in a Link GAT-1 flight simulator at three alcohol dose levels. Aerospace Medicine. 45, 33- 44.

Jones, H. and Higdon, A. 1991. HAWK data acquisition system users' guide. Unpublished internal working document, U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL.

Rehmann, J.T. 1982. Pilot performance measurement: An annotated bibliography. Tech. Report DOT/FAA/EM-81/16. Atlantic City: FAA Technical Center.

Simmons, R. R., Caldwell, J. A., Stephens, R. L., Stone, L. W., Carter, D. J., Behar, I., Mitchell, G. W., Knox, F. S., Jones, H. D., and Taylor, P. L. 1989. Effects of the chemical defense antidote atropine sulfate on helicopter pilot performance: A simulator study. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 89-17.

Stephens, R. L., Caldwell, J. A., Comperatore, C. A., Pearson, J. Y., and Delrie, D. M. 1992. Effects of terfenadine and diphenhydramine on brain activity and performance in a UH-60 flight simulator. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Report No. 92-33.